

**DEPARTMENT OF VETERANS AFFAIRS**

**Office of Informatics and Analytics**

**Office of Knowledge Based Systems (KBS)**

**Informatics Architecture Support Services**

**Contract No. VA701-16-C-0157**

**Obligation No. 776-C60159**

**CLIN: 1017**

***SUPPORTING DOCUMENT***

**Task 5.5 Identify SOLOR Content that Requires Special Handling**

***Version 1.0***

**Date:** **January 31, 2019**

Table of Contents

[1. Purpose 3](#_Toc533089827)

[2. Approach to Identify Content 3](#_Toc533089828)

[3. Rules for Evaluating Membership in RefSets 6](#_Toc533089829)

[4. Rules for placing concepts in the RefSets 7](#_Toc533089830)

[5. Other Symmetry Issues 8](#_Toc533089831)

[6. Grades, Scales and Scores 8](#_Toc533089832)

# Purpose

The purpose of the RefSets produced in the four-month extension to Option Year 1 was to identify concepts in SNOMED that are and are not symmetrical.

# Approach to Identify Content

We consider modeling “symmetrical” if concepts, which are opposites of each other (inverse concepts)

* exist in SNOMED
* reside in the correct hierarchy under the correct parent concept

Example:

299331007 |Knee joint varus deformity (finding)| has two children, which are opposites. Both are present and under the correct parent concept:

64925008 |Acquired genu varum (disorder)|

79168008 |Congenital genu varum (disorder)|

Inverse concepts do not necessarily have to reside under the same parent to be considered symmetrically modeled.

Example:

230763008 |Traumatic cerebral edema (disorder)| and 330011000119102 |Non-traumatic cerebral edema (disorder)| are inverse.

230763008 |Traumatic cerebral edema (disorder)| is a child of 127295002 |Traumatic brain injury (disorder)|

330011000119102 |Non-traumatic cerebral edema (disorder)| is a child of 2032001 |Cerebral edema (disorder)|

The below approach was used to identify the first RefSet, which included querying 50,000 concepts:

1. Missing Content – Via Inverse Work
	1. Inverse Refset work has identified roughly 6,000 concepts that need to be reviewed to confirm missing opposing concepts

Table 1. Example of missing opposing concepts

|  |  |
| --- | --- |
| Conceptid1 | FSN |
| 8587003 | Congenital diverticulum of colon (disorder)Missing opposite: Acquired diverticulum of colon (disorder) |
| 8656007 | Total traumatic cataract (disorder)Missing opposite: Partial traumatic cataract (disorder) |
| 9027003 | Normal pulmonary arterial wedge pressure (finding)Missing opposite: Abnormal pulmonary arterial wedge pressure (finding) |
| 21370008 | Tenotomy of abductor of hip, open (procedure)Missing opposite: Tenotomy of abductor of hip, closed (procedure) |

1. Missing Content – Via Leaf Nodes
2. Identify all concepts that are parents of a leaf with only one leaf.
3. Content Modeled Inappropriately:
	1. Concepts that are inferred with more than one of the same Attribute Type
	2. Remove any Concepts that are modeled with more than one of the same Attribute Type and the same Value
	3. Remove any Concepts with Attributes that are frequently used with different values, like Finding Site or Associated Morphology
	4. Remove any Concepts from hierarchies that will not be reviewed (Products, Substances, Qualifier value, Situations, Body structures)

Figure 1: Content Modeled Inappropriately



1. Concept Modelled Inappropriately – Inverse
	1. Using concepts that are paired as inverse of each other and identify those that are inappropriately modeled based on radical modeling differences.

Figure 2. Example of Inverse Concepts modeled with radical differences



1. Identify concepts that contain a common phrase without appropriate role
	1. Find all concepts that have common phrases like “Acute”, “Chronic”, “Acquired”, “Congenital” that do not have the corresponding attribute.

Figure 3. FSN contains "Acute", but does not have a Clinical Course = Acute



1. Grades, Stages, Scales, and Scores
	1. Review concepts, that represent Grades, Stages, Scales, and Scores to ensure all are present in the Finding and Disorder hierarchies.

# Rules for Evaluating Membership in RefSets

For this project, we will deliver four RefSets that will categorize our efforts as follows:

1. Symmetric Correct Modeling
	* A simple RefSet of child concepts that were reviewed and deemed to be in the correct hierarchy and under the correct parent.
2. Symmetric Incorrect Modeling
	* A simple RefSet of child concepts that were reviewed and deemed to be placed in the wrong hierarchy (under an incorrect parent).
3. Symmetric Modeling Children Present
	* A simple RefSet of parent concepts that had correct symmetric children
4. Symmetric Modeling Non-existent Children
	* An Annotation RefSet with parent concepts that are missing symmetric children that should exist and any comments on what needs to be done to make them symmetric.

# Rules for placing concepts in the RefSets

* Inverse concepts
	+ If an inverse concept has an existing opposite concept and it is in the appropriate hierarchy, it was considered Symmetric Correct Modelling.
	+ If an inverse concept has an existing opposite concept and it is in the wrong hierarchy, it was considered Symmetric Incorrect Modelling.
* Parents of leaf concepts (concepts with only one child) were reviewed for:
	+ If the child is in the correct hierarchy it was placed in
	+ Is the child modeled correctly
		- “correct modeling” only applies to the correct inferred view for this concept as it pertains to symmetry
		- If a concept has other modeling problems, it is not marked as “incorrectly modelled”.
* Some keywords that could indicate the need for symmetry are not always reliable, for example:
	+ *Traumatic vs. non-traumatic* - concepts without a stated “traumatic” in the FSN are considered non-traumatic by default.
	+ *With vs. without* - not every concept that has a “with” or “without” needs its opposite, e.g. Diagnostic arthroscopy of elbow with synovial biopsy (procedure) does not need a “…without biopsy”.

# Other Symmetry Issues

* Clinical Course vs. Associated Morphology

Throughout SNOMED, inconsistent modeling using attributes “clinical course” and “associated morphology” exists.

Example:

19429009 |Chronic ulcer of skin (disorder)| is modelled using 116676008 |Associated morphology (attribute)| = 405719001 |Chronic ulcer (morphologic abnormality)|

111422001 |Chronic abscess of breast (disorder)| is modelled using both the |Associated morphology (attribute)| = 79203009 |Chronic abscess (morphologic abnormality)| and the 263502005 |Clinical course (attribute)| = 90734009 |Chronic (qualifier value)|

This inconsistency also leads to “asymmetric” modeling, but was not in scope for this task.

# Grades, Scales and Scores

This following analysis of the inconsistent use of Procedures and/or Observables as the value of the “Interprets” Attribute is exploratory and not part of the RefSet creation.

The Findings and Disorders reviewed were found to use a Procedure 42 times vs. and Observable Entity 352 times. In 41 cases, both a Procedure and Observable Entity were used for the Interprets attribute. 400 of the concepts had no Interprets Attribute at all.

Figure 4: Grade concept with an Interprets = Procedure



Figure 5: Grade concept with an Interprets = Observable Entity



Figure 6: Grade Concept with both a Procedure and Observable used for the Interprets Attribute



Figure 7: Grade with no Interprets Attribute



Potential changes to Grades, Scales and Scores Concepts

For “symmetry” reasons:

* Each finding with “grade, score, stage etc.” should be consistently modeled with an Interprets = Observable Entity, which would be the score or grade.

Example:

 385377005 |Gleason grade finding for prostatic cancer (finding)|:
        116676008 |Associated morphology (attribute)| = 52988006 |Lesion (morphologic abnormality)|,
        363714003 |Interprets (attribute)| = 372278000 |Gleason score (observable entity)|

* If the Observable Entity concept does not exist, it should be created.
* Another option would be:
	+ Each finding should be modeled with Interprets = Assessment using…scale/score….(procedure)”
	+ The Assessment using…scale/score….(procedure) should be modeled with Method = 129268004 |Observation - action (qualifier value)| and Has focus = “whatever the focus of the score is”
		- In case of the Gleason score, it would be 254900004 |Carcinoma of prostate (disorder)|
	+ However, there are significantly less Procedure concepts and would require a lot of additions to SNOMED.

# Appendix

Table of Figures

[Figure 1: Content Modeled Inappropriately 5](#_Toc533090145)

[Figure 2. Example of Inverse Concepts modeled with radical differences 6](#_Toc533090146)

[Figure 3. FSN contains "Acute", but does not have a Clinical Course = Acute 6](#_Toc533090147)

[Figure 4: Grade concept with an Interprets = Procedure 8](#_Toc533090148)

[Figure 5: Grade concept with an Interprets = Observable Entity 9](#_Toc533090149)

[Figure 6: Grade Concept with both a Procedure and Observable used for the Interprets Attribute 9](#_Toc533090150)

[Figure 7: Grade with no Interprets Attribute 9](#_Toc533090151)